The LOFAR Epoch of Reionization Data Model

Lampropoulos, Panagiotis

Appendix A

Polarimetric issues

In this appendix we introduce several fundamental concepts in the description of polarized radiation. We feel that they are important to mention and will be used in forthcoming publications to guide us to a better understanding of complex calibration processes in radio interferometry.

A.1 The Electric-Field Vector and Coherency Matrix

The effects of linear passive media on the propagated photons can be represented by linear transformations of the electric field variables. The nature of those effects, the spectral profile of the light and the chromatic and polarizing properties of the medium through which light passes, all affect the degree of mutual coherence. In general coherent interactions can be represented by the Jones calculus (Jones, 1948, 1942), while incoherent interactions of polychromatic light require the Mueller calculus (Barakat, 1963), since the loss of coherence needs more parameters to be described. The two components of the electric field (e.g. those received at two dipoles) can be arranged as the components a 2×1 complex vector:

$$\mathbf{e}(t) = \begin{pmatrix} E_x(t) \\ E_y(t)\, e^{i\delta(t)} \end{pmatrix} \tag{A.1}$$

where $\delta(t)$ is the relative phase. This vector includes all information about the temporal evolution of the electric field. When the parameters have no time dependence this is called the *Jones vector*. Moreover, the coherency (or polarization or density) matrix of a light beam contains all the information about its polarization state. This Hermitian 2×2 matrix is defined as

$$\mathbf{C} \equiv \left\langle \mathbf{e}(t) \otimes \mathbf{e}^{\dagger}(t) \right\rangle = \begin{pmatrix} \langle e_1(t)\, e_1^*(t) \rangle & \langle e_1(t)\, e_2^*(t) \rangle \\ \langle e_2(t)\, e_1^*(t) \rangle & \langle e_2(t)\, e_2^*(t) \rangle \end{pmatrix} \tag{A.2}$$

This is the coherency matrix of the perpendicular dipoles of a single LOFAR HBA antenna. $\otimes$ stands for the Kronecker product and the brackets indicate averaging over time (Boonstra, 2005). The coherency matrix is a correlation matrix whose elements are the second moments of the signal. Using the ergodic hypothesis the brackets can be considered as ensemble averaging. Due to its statistical nature its eigenvalues ought to be non-negative. The normalized version of this matrix $\hat{\mathbf{C}} = \frac{\mathbf{C}}{\mathrm{tr}(\mathbf{C})}$ contains information about the population and coherencies of the polarization states (Fano, 1957). This object is the equivalent of the single brightness point in the scalar version of the theory.

The measurable quantities Stokes I, Q, U and V arise as the coefficients of the projection of the coherency matrix onto a set of Hermitian trace-orthogonal matrices, the generators of the unitary SU(2) group plus the identity matrix. Parameters with direct physical meaning can be derived from the corresponding measurable quantities. The Stokes parameters are usually arranged as a 4×1 vector,

$$\mathbf{s} = \begin{pmatrix} I \\ Q \\ U \\ V \end{pmatrix}.$$

An alternative notation is the 2×2 Stokes matrix:

$$\mathbf{S} = \frac{1}{2} \begin{pmatrix} I+Q & U-iV \\ U+iV & I-Q \end{pmatrix} \equiv \mathbf{C}, \tag{A.3}$$

which relates the measured coherency matrix quantities to the Stokes parameters (Born & Wolf, 1999).

A.2 The Jones Formalism

An adequate method to describe a non-depolarizing system is the Jones formalism. It represents the effects on the polarization properties of an EM wave after the interaction with such a system. For passive, pure systems, the electric field components of the light interacting with them is given by the corresponding Jones matrix $\mathbf{J}$,

$$\mathbf{e}' = \mathbf{J}\mathbf{e}.$$

As both the initial and final fields can fluctuate, it is useful to describe the properties of partially polarized light with the coherency matrix. Thus,

$$\begin{aligned} \mathbf{C}' &= \left\langle \mathbf{e}' \otimes \mathbf{e}'^{\dagger} \right\rangle = \left\langle (\mathbf{J}\mathbf{e}) \otimes (\mathbf{J}\mathbf{e})^{\dagger} \right\rangle \\ &= \left\langle \mathbf{J}\mathbf{e} \otimes \mathbf{e}^{\dagger}\mathbf{J}^{\dagger} \right\rangle = \mathbf{J} \left\langle \mathbf{e} \otimes \mathbf{e}^{\dagger} \right\rangle \mathbf{J}^{\dagger} = \mathbf{J}\mathbf{C}\mathbf{J}^{\dagger} \end{aligned} \tag{A.4}$$

As we are dealing with interferometry, the two $\mathbf{J}$ matrices can come from two different telescopes. The effects on the electric field vectors in the coherency matrix $\mathbf{C}'$ can be written as an operation of these Jones matrices on the original unaffected coherency matrix $\mathbf{C}$. As we already mentioned, the coherency matrix can also be written as a four vector with $\mathbf{c} = \left(\langle e_1 e_1^* \rangle, \langle e_1 e_2^* \rangle, \langle e_2 e_1^* \rangle, \langle e_2 e_2^* \rangle \right)$. This vector is related to the Stokes vector via

$$\mathbf{s} = \mathbf{L}\mathbf{c} \quad \text{with} \quad \mathbf{L} = \begin{pmatrix} 1 & 0 & 0 & 1 \\ 1 & 0 & 0 & -1 \\ 0 & 1 & 1 & 0 \\ 0 & -i & i & 0 \end{pmatrix}. \tag{A.5}$$

The matrix has the following property $\mathbf{L}^{-1} = {}^1\!/_2\mathbf{L}^\dagger$. Using the properties of the Kronecker product we then find, in terms of Stokes parameters, that

$$\begin{aligned} \mathbf{s}' = \mathbf{L}\left\langle \mathbf{Je} \otimes (\mathbf{Je})^\dagger \right\rangle = \mathbf{Ns} \\ \mathbf{N} = \mathbf{L}\left(\mathbf{J} \otimes \mathbf{J}\right)\mathbf{L}^{-1} \\ \text{with} \\ \mathbf{N}_{kl} = \frac{1}{2}\mathrm{tr}\left(\sigma_k \mathbf{J} \sigma_l \mathbf{J}^\dagger\right). \end{aligned} \tag{A.6}$$

where σ_i are the Pauli matrices. A Jones matrix can represent a physically realizable state as long as the transmittance condition (gain or intensity transmittance) holds; that is, the ratio of the initial and final intensities must be $0 \leqslant g \leqslant 1$. The reciprocity condition describes the effect when the output signal follows the path in the inverse order. For every proper Jones matrix $\mathbf{e}' = \mathbf{J}^\dagger \mathbf{e}$. This result does not hold when magneto-optic effects are present. In this case the Mueller–Jones matrices have to be used. If a Jones matrix represents a physically realizable state the reciprocal matrix also represents a physical effect.

A.3 Pauli Matrices

Pauli matrices are well known and have been used for the analysis of partially polarized light (Fano, 1954). Their major advantage is that they satisfy a set of properties that significantly reduce the complexity of calculations associated with the intensity. The identity plus the Pauli matrices in two dimensions are defined as

$$\begin{aligned} \sigma_0 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \quad \sigma_1 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \\ \sigma_2 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \quad \sigma_3 = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} \end{aligned} \tag{A.7}$$

This set of 2×2 linearly independent matrices constitute a basis for the vector space of 2×2 Hermitian matrices over the complex numbers. Summarizing their properties, they are Hermitian and they follow the commutation relations $\lfloor \sigma_i, \sigma_j \rfloor = \sigma_i \sigma_j - \sigma_j \sigma_i = i2\varepsilon_{ijk}\sigma_k$ where ε_{ijk} is the Levi-Civita permutation symbol. These matrices are unitary and traceless except for the identity matrix. The linear expansion of the coherency matrix in this basis is

$$\mathbf{C} = \frac{1}{2}\sum \mathrm{tr}\left(\mathbf{C}\sigma_i\right)\sigma_i \tag{A.8}$$

with $s_i = \mathrm{tr}\left(\mathbf{C}\sigma_i\right)$ being the four Stokes parameters: $i = 0, 1, 2, 3$ corresponding to the Stokes I, Q, U and V, respectively.

A.4 The Stokes Vector and Matrix

In the literature the Stokes parameters are usually arranged as a 4×1 vector, $\mathbf{s} = (I, Q, U, V)^T$ (Gil, 2007). An alternative notation would be to introduce a 2×2 Stokes matrix.

$$\mathbf{S} = \frac{1}{2} \begin{pmatrix} I+Q & U-iV \\ U+iV & I-Q \end{pmatrix} \tag{A.9}$$

In this case for pure states we have $\|\mathbf{C}\|^2 = \frac{1}{2}\|\mathbf{S}\|^2$ and in the case of unpolarized light $\|\mathbf{C}\|^2 = \frac{1}{4}\|\mathbf{S}\|^2$. This is because for unpolarized light Q, U and V are zero and the matrix becomes I multiplied by the identity matrix. Since there is a factor of $\frac{1}{2}$, the Frobenius norm of the matrix has this factor squared. Even if the source is not polarized one should observe the same signal in both orthogonal polarizations, as in this case both polarization states are equiprobable. This is an important fact for calibration. The X and Y dipoles of the LOFAR HBAs should measure the same signal, for unpolarized sources, and any fluctuations should be due to polarization calibration errors.

A.5 Polarization Level

In addition, the degree of polarization is defined as $P \equiv \sqrt{U^2 + V^2 + Q^2}/I$. We can also define the vector (Gil, 2007)

$$\mathbf{p}^T = \frac{1}{PI} \begin{pmatrix} Q \\ U \\ V \end{pmatrix}$$

The Stokes vector can be decomposed using those parameters in several ways. A trivial decomposition is between a polarized and unpolarized state. As radio receivers are intrinsically polarized, and in the case of LOFAR orthogonal, a spectral decomposition can express the Stokes vector as a convex linear sum of two orthogonal states. This makes the relationship between the measured signals and the parameters describing the system more clear, since we have to deal with orthogonal states. The spectral decomposition of the coherency matrix is using its eigenvalue structure to decompose it into pure states, i.e. eigenvectors. The spectral decomposition (diagonalization) of the coherency matrix is then equivalent to

$$\mathbf{s} = I \times \left(\frac{1+P}{2} \begin{bmatrix} 1 \\ \mathbf{p}^T \end{bmatrix} + \frac{1-P}{2} \begin{bmatrix} 1 \\ -\mathbf{p}^T \end{bmatrix} \right)$$

Of course, for a mixed state there are infinite combinations of independent states into which it can be decomposed, but it would be useful if the selection is such that it matches the characteristics of the system. The parameters I and P are invariant under unitary transformations (changes of coordinate systems). They are directly related to the eigenvalues of $\mathbf{C}$. Pure states are related to rank-1 polarization matrices, and mixed states to rank-2. Wolf showed that there always exist two orthogonal reference directions, such that the degree of coherence is maximized and coincides with P. This is important because various random distributions correspond to unpolarized light. The measurement of the correlations of the Stokes parameters allows us to distinguish between those different types of non-polarized light.

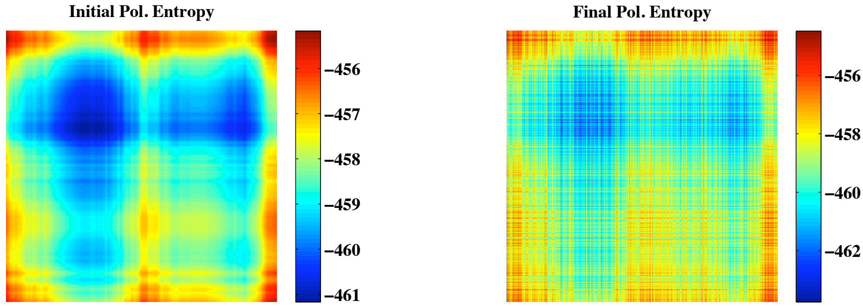

Figure A.1: Left: the polarization entropy of a simulated Galactic diffuse synchrotron emission map. Right: the same plot after applying a depolarizing effect (Faraday rotation) along the horizontal axis

A.6 Entropy

Concluding with the quality criteria, the von Neumann entropy can be applied to electromagnetic waves (Fano, 1957). In terms of the coherency matrix it is defined as $S = -\mathrm{tr}\left(\hat{\mathbf{C}} \ln \hat{\mathbf{C}}\right)$. Is a measure of the difference in the amount of information between pure and mixed states, both with same intensity. It can be expressed as a function of the eigenvalues of $\mathbf{C}$ as $S = -\frac{1}{2}\left[(1+P)\ln\left(\sqrt{1+P}\right) + (1-P)\ln\left(\sqrt{1-P}\right)\right]$. It is a decreasing monotonic, bounded function of P (Gil, 2007). It attains its maximum value, $S = \ln(2)$, for non polarized light and its minimum, $S = 0$, for totally polarized light. Using this entropy one can define the polarization temperature. Depolarizing effects, such as the ionospheric Faraday rotation, can be studied this way in order to detect spatial heterogeneity. For example, ionospheric TIDs should lead to an increase in the polarization entropy. This is a good scheme of ranking observations: if the polarization entropy differs between two maps, residual Faraday rotation and leakage can be present. Finally, for non-Gaussian distributions of polarization states, higher order moments are needed and the Stokes system is no longer adequate to describe the polarization states.

The standard (self)-calibration procedure for interferometric observation of polarized light aims at recovering the coherency matrix of the polarized radiation. Given the statistical nature of the coherency matrix, we must emphasize the importance of parameters that give a measurement of their polarimetric purity. Polarization entropy is a concept related to the impurities of the media through which radiation propagates, as it was defined above. It is useful for many purposes, especially when depolarization is a relevant subject. An increase in the polarization entropy signifies a decrease in the polarization purity. It is a direct way to access the quality of the data. Ionospheric Faraday rotation causes a depolarization with a certain frequency behavior. We would expect that the entropy would follow this behavior. In Figure A.1 we show the polarization entropy for several lines of sight as a function of frequency in a simulated map. Attention should be brought to the fact that is really hard to distinguish between Faraday rotation and other depolarization effects.

A.7 Lorentz Transformations

The transformations of the Stokes vector lead to a system of differential equations. From the statistical interpretation of the coherency matrix we derive that the Stokes I parameter has to be positive and that $s_0 = I \geq s_1^2 + s_2^2 + s_3^2 = Q^2 + U^2 + V^2$, which means that there can be no more polarized light than the total light. This is a key concept in physics, namely the conservation of energy. Extending this remark, we can define a Minkowski space (in the mathematical sense) for the Stokes 4-vectors. The norm in this space would be $\|\mathbf{s}\| = I^2 - Q^2 + U^2 + V^2$. Positive values of the norm correspond to the time-like vectors of the special theory of relativity. Light-like Stokes vectors correspond to totally polarized states. The Poincaré sphere defines the light cone with the exception that the symmetric part of the cone does not have any special meaning. The Jones vector transforms as usual under the Lorentz group. The coherency matrix is defined through the Kronecker product of two Jones vectors and is a 2×2 spinor of rank two. The Stokes vector and the coherency matrix must be realizations of the same irreducible representation of the Lorentz group, as they both have 4 independent components. The Lorentz transformations form a 6 parameter group. The six generators are the 3 spatial rotation generators of O(3) $\mathbf{R}$ and the Lorentz boosts $\mathbf{B}$. The infinitesimal transformation of the Stokes vector is

$$\mathbf{s}' = \mathbf{s} - \sum_{i}^{3} (r_i \mathbf{R}_i + b_i \mathbf{B}_i) \mathbf{s} \, \mathrm{d}\, I = \mathbf{s} - \mathbf{K}\mathbf{s} \, \mathrm{d}\, I$$

where $\mathbf{K}$ is a matrix that resembles the absorption matrix. This equation is the radiative transfer equation along the signal path. If one choses to exclude effects that intensify the radiation, this is the complete mathematical description of the effects in the signal path. The Lorentz boosts describe polarizing effects, while the spatial rotations describe the Faraday rotation. The relation between the initial and final Stokes vector is a finite Lorentz transformation. While it sounds straightforward it is not an easy task. The problem arises from the non commutativity of the generators. Magnus has proposed a solution to this problem. We will discuss this soon.

A.8 Clifford Algebra

After having introduced the coherency matrix and the Jones formalism we proceed a step further in the mathematical abstraction in order to unify all possible cases. We note that every system is equivalent to a parallel combination of pure systems, and any pure system to a combination of retarders and de-attenuators. In Hamaker (2000) the author proposes the use of the quaternion algebra to describe the Jones matrices, or their $SO^+(1,3)$ covering group. Quaternions, as any other space isomorphic to $\mathbb{C}$ constitute a sub-algebra of a Clifford algebra. In particular, in three dimensions the Clifford product is equivalent to spatial rotations. Pure systems can still be described as Lorentz transformations in this algebra. Since algebraic manipulations become more clear, the computational requirements of problems involving partial polarization can be reduced. The Clifford algebra of three-dimensional space Cl_3 represents the four dimensional Minkowski space time. Clifford algebras have been used extensively to describe polarization mode dispersion (PMD) in optical fibers (Reimer et al., 2008). In that field they have to deal with a spatially inhomogeneous, birefringent medium (the optical fibre), which resembles the effects caused by the ionosphere on the EM waves that propagate through it.

A.9 Magnus Expansion

In the case of polarization mode dispersion, which can occur due to ionospheric Faraday rotation, transversal of the polarized source signal through a set of phase screens etc, one is interested in recovering the original polarization state of the radiation. Two possibly quasi-orthogonal modes, represented by the input frequency-independent and the output Jones vectors, are related through a complex 2×2 Jones matrix. Despite the fact that we might not have any prior knowledge about the structure of the medium through which radiation propagates, the frequency dependence of that effect contains useful information about that medium. For example the $\sim \lambda^2$ Faraday rotation should leave a distinct imprint on the signal, which can help us distinguish it from other depolarizing effects. In the general case, the coherency matrix is transformed as $\mathbf{C}' = \mathbf{JCJ}^{\dagger}$. An arbitrary Jones matrix with determinant 1, can be also expressed in terms of two vectors $\mathbf{a}$ and $\mathbf{b}$ according to

$$\mathbf{J} = \exp\left[-\left(i/2\right)\left(\mathbf{b} + i\mathbf{a}\right) \cdot \sigma\right], \tag{A.10}$$

where σ is the Pauli spin vector. This is a Jones matrix representation of the Lorentz transformation in a Clifford algebra. In the case of a frequency dependent effect, $\mathbf{J}\left(\omega\right)$, the Jones space operator can be decomposed into real and imaginary components as

$$\begin{aligned} \mathbf{J}\left(\omega\right) &= \exp\left[-\left(i/2\right)\left(\mathbf{b} + i\mathbf{a}\right) \cdot \sigma\right], \\ \tilde{\mathbf{J}}_!\tilde{\mathbf{J}} &= -\frac{i}{2}\left[\boldsymbol{\Omega} + i\boldsymbol{\Lambda}\right], \end{aligned} \tag{A.11}$$

where the subscript ω represents differentiation with respect to the frequency. A solution for $\mathbf{J}$ can be obtained through the Magnus expansion, which specifies $\mathbf{J}\left(\omega\right) = \exp\left(\sum_{n=0}^{\infty} \mathbf{B}_n\left(\omega\right)\right) \mathbf{J}\left(\omega_0\right)$, where the first two coefficients are given by

$$\mathbf{B}_1\left(\omega\right) = \int_{\omega_0}^{\omega} d\omega_1 J_\omega\left(\omega_1\right),$$
$$\mathbf{B}_2\left(\omega\right) = \int_{\omega_0}^{\omega}\int_{\omega_0}^{\omega_1} d\omega_2 d\omega_1\left(J_\omega\left(\omega_1\right) J_\omega\left(\omega_2\right) - J_\omega\left(\omega_2\right) J_\omega\left(\omega_1\right)\right).$$

The coefficients of the Magnus expansion for $n > 2$, are related to those of lower order through recursion. Taylor expanding the frequency derivatives of the Jones matrix to third order gives us a way to directly evaluate the Magnus coefficients. To the extent of our knowledge, the methods mentioned above have not been applied in the field of radio interferometry.

Appendix B

Mathematical definitions

B.1 Imaging with non-coplanar arrays

The measurement equation for a non-coplanar array can be written as:

$$\mathbf{R} = \mathbf{A}\mathbf{C}\mathbf{A}^H,$$

where

$$\begin{aligned}
\mathbf{R} &= \mathbf{R}(t) \\
\mathbf{A} &= \mathbf{A}(t) = \left[\mathbf{a}\left(l_1, m_1\right), \cdots, \mathbf{a}\left(l_n, m_n\right)\right] \\
\mathbf{a}\left(l_n, m_n\right) &= \begin{bmatrix} e^{-2\pi i\left(u_{10}l + v_{10}m + w_{10}n\right)} \\ \vdots \\ e^{-2\pi i\left(u_{1k}l + v_{1k}m + w_{1k}n\right)} \end{bmatrix} \\
\mathbf{C} &= diag\left[\begin{array}{ccc} \frac{I(l_1, m_1)}{\sqrt{1-l_1^2-m_1^2}}, & \cdots, & \frac{I(l_n, m_n)}{\sqrt{1-l_n^2-m_n^2}} \end{array}\right]
\end{aligned} .$$

Then the following equation has to be inverted to solve the imaging problem:

$$\mathbf{R} = \mathbf{G}\mathbf{A}(l, m)\mathbf{C}\mathbf{A}^H(l, m)\mathbf{G}^H + \mathbf{N},$$

where **N** is the covariance matrix of additive white noise.

B.2 Polarimetric imaging

LOFAR antennas receive two orthogonal polarization (along the X and Y axis of the antennas). The array response vector is now replaced by two orthogonal vectors describing the two linear polarization states so that $\mathbf{a}(l, m)_X \perp \mathbf{a}(l, m)_Y$. The polarization state of a monochromatic source is described by the coherence matrix.

Then the matrix form of the calibrated ME becomes:

$$
\begin{aligned}
&\mathbf{R} = \mathbf{G}\mathbf{A}(l,m)\mathbf{C}\mathbf{A}^H(l,m)\mathbf{G}^H + \mathbf{N} \\
&\mathbf{A} = \mathbf{A}(t) = \left[\mathbf{a}\left(l_1,m_1\right)_X, \mathbf{a}\left(l_1,m_1\right)_Y, \cdots, \mathbf{a}\left(l_n,m_n\right)_X, \mathbf{a}\left(l_n,m_n\right)_Y\right] \\
&\mathbf{C} = \begin{bmatrix} \begin{matrix} C_{11}^1 & C_{12}^1 \\ C_{21}^1 & C_{22}^1 \end{matrix} & & \\ & \ddots & \\ & & \begin{matrix} C_{11}^n & C_{12}^n \\ C_{21}^n & C_{22}^n \end{matrix} \end{bmatrix} \\
&\mathbf{G} = diag\left(\mathbf{G}_1, \cdots, \mathbf{G}_p\right) \\
&\mathbf{G}_i = \begin{bmatrix} g_{11,i} & g_{12,i} \\ g_{21,i} & g_{22,i} \end{bmatrix}
\end{aligned}
$$